Вычислительно трудные задачи и дерандомизация
Лекция 5': Псевдо-случайный генератор Нисана-Вигдерсона (окончание)

Дмитрий Ицыксон

ПОМИ РАН

12 апреля 2009

Псевдослучайный генератор

Псевдослучайное распределение

Распределение R на $\{0,1\}^m$ называется (S,ϵ)-псевдослучайным, если для каждой схемы C размера $\leq S$:

$$|\Pr[C(R)=1]-\Pr[C(U_m)=1]|<\epsilon$$

Псевдослучайный генератор

$S(\ell)$ — правильная неубывающая функция. Функция $G:\{0,1\}^*\to\{0,1\}^*$ называется $S(\ell)$-псевдослучайным генератором, если

- G вычислима за время $2^{O(n)}$
- $\forall z\in\{0,1\}^\ell$, $|G(Z)|=S(\ell)$
- $\forall\ell$ распределение $G(U_l)$ является $(S(\ell)^3,\frac{1}{10})$-псевдослучайным.

Чем помогает псевдослучайный генератор?

- Если существует $2^{\epsilon\ell}$-псевдослучайный генератор ($\epsilon > 0$), то $\mathbf{BPP} = \mathbf{P}$.
- Если существует 2^{ℓ^ϵ}-псевдослучайный генератор ($\epsilon > 0$), то $\mathbf{BPP} \subseteq \mathbf{QuasiP} = \mathbf{DTime}[2^{polylog(n)}]$.
- Если существует $\ell^{\omega(1)}$-псевдослучайный генератор то $\mathbf{BPP} \subseteq \mathbf{SUBEXP} = \cap_{\epsilon>0}\mathbf{DTime}[2^{n^\epsilon}]$.

Схемная сложность функций

$f : \{0,1\}^n \to \{0,1\}$.

Сложность в наихудшем случае

$$H_{wrs}(f) = \max\{S | \forall \text{ схемы } C \text{ размера } \leq S \; \exists x \in \{0,1\}^n : C(x) \neq f(x)\}$$

Сложность в среднем случае

$$H_{avg}(f) = \max\{S | \forall \text{ схемы } C \text{ размера } \leq S \; \Pr_{x \leftarrow U(\{0,1\}^n)}[C(x) = f(x)] < \frac{1}{2} + \frac{1}{S}\}$$

Генератор из сложных функций

Теорема. (Нисан-Вигдерсон) Если существует $f : \{0,1\}^* \to \{0,1\}$, вычислимая за время $2^{O(n)}$, что

- $H_{avg}(f) \geq 2^{\epsilon n}$, тогда существует $2^{\epsilon' n}$-псевдослучайный генератор.
- $H_{avg}(f) \geq 2^{n^{\epsilon}}$, тогда существует $2^{n^{\epsilon'}}$-псевдослучайный генератор.
- $H_{avg}(f) \geq n^{\omega(1)}$, тогда существует $n^{\omega(1)}$-псевдослучайный генератор.

Генератор Нисана-Вигдерсона

Определение. (Комбинаторный дизайн) Семейство $\mathcal{I} = \{I_1, I_2, \ldots, I_m\}$ подмножеств $\{1, 2, \ldots, \ell\}$ называется (ℓ, d, n)-дизайном, если $\forall j, |I_j| = n$ и $\forall j \neq k, |I_j \cap I_k| \leq d$.

Определение. (Генератор Нисана-Вигдерсона) Пусть $\mathcal{I} = \{I_1, I_2, \ldots, I_m\}$ — дизайн. $f : \{0,1\}^* \to \{0,1\}$. $NW^f_{\mathcal{I}} : \{0,1\}^\ell \to \{0,1\}^m$:

$$NW^f_{\mathcal{I}}(z) = f(z_{I_1}) \circ f(z_{I_2}) \circ \cdots \circ f(z_{I_m})$$

Лемма. (Конструкция дизайна) За время $2^{O(\ell)}$ можно построить (ℓ, d, n)-дизайн, где $n > d, \ell > 10n^2/d$. В этом дизайне будет $2^{d/10}$ множеств.

Оценки Чернова

Теорема. $X_i \in \{0, 1\}$ — независимые одинаково распределенные случайные величины. $\mathrm{E} \sum_{i=1}^{n} X_i = \mu$. Тогда

$$\Pr\left[\sum_{i=1}^{n} X_i > (1+\varepsilon)\mu\right] \leq \left(\frac{e^{\varepsilon}}{(1+\varepsilon)^{(1+\varepsilon)}}\right)^{\mu}$$

$$\Pr\left[\sum_{i=1}^{n} X_i < (1-\varepsilon)\mu\right] \leq e^{-\epsilon^2\mu/2}$$

Построение дизайна

Лемма. (Констуркция дизайна) За время $2^{O(\ell)}$ можно построить (ℓ, d, n)-дизайн, для любых $n > d, \ell > 10n^2/d$. В этом дизайне будет $2^{d/10}$ множеств.

Доказательство.

- Можно считать, что $\ell < 20n^2/d$. Добавлять n-элементные множества по одному, пока их меньше 2^d, так чтобы все пересечения не превосходили d.
- Случайное множество I: каждый элемент включаем с вероятностью $\frac{2n}{\ell}$.
- $|I| = \sum_{j=1}^{l} x_j$, $\mathsf{E}\,|I| = \sum_{j=1}^{l} \mathsf{E}\,x_i = 2n$, отсюда $\Pr[|I| \geq n] \geq 0.9$
- $d/10 < \mathsf{E}\,|I \cap I_k| = 2n^2/\ell < d/5$, то $\Pr[|I \cap I_k| \geq d] \leq \left(\frac{e^4}{5^5}\right)^{d/10} < 0.5 \cdot 2^{-d/10}$
- $\Pr[\forall j |I \cap I_j| \geq d] \leq 0.5$
- $\Pr[|I| \geq n \wedge \forall j |I \cap I_j| \geq d] \geq 0.4$
- Осталось выкинуть из I лишние элементы.